The Effect of Non Gaussian Errors on the Determination of Steeply Falling Spectra

C.O. Escobar[1], L. G. dos Santos[1], and R.A. Vázquez[2]

[1] *Instituto de Fisica Gleb Wataghim, Unicamp*
[2] *Departamento de Física de Partículas,*
Universidade de Santiago, 15706 Santiago de Compostela, Spain

Abstract

The determination of a steeply falling energy spectrum from rare events with non-gaussian measurement errors is a delicate matter. The final shape of the spectrum may be severely distorted as a consequence of non-gaussian tails in the energy resolution of the experiment. We illustrate this effect with the recent experimental efforts to determine the ultra-high energy extreme of the cosmic ray spectrum.

1 Introduction

We will study the determination of a steeply falling spectrum when the energy measurement errors show long, asymmetrical, non-gaussian tails. We will see that for a steep spectrum the effect of long tail error functions can be dramatic: spectrum slopes, cutoff or other features could be changed or even erased, depending on the degree of deviation of the error function (energy resolution) from a gaussian distribution.

This problem, well known to nuclear spectroscopists working near the end point of a spectrum [1] is, however, often ignored in areas such as Cosmic Ray physics and we therefore choose to illustrate it by examining the spectrum of Ultra High Energy Cosmic Rays (UHECR) as determined by ground arrays, addressing the crucial question of the existence or absence of a cutoff in the spectrum, known as the Greisen-Zatsepin-Kuzmin (GZK)cutoff [2].

2 Spectrum reconstruction with non-gaussian errors

We will consider four cases for the reconstruction error function. These cases illustrate the different possibilities.

Let's start with the simplest, well-known, case of a gaussian error reconstruction function. Assume that events of energy E' are reconstructed with energy E with a probability given by

$$P(E, E') = N \exp(\frac{-(E - E')^2}{2\sigma^2}),$$ (1)

where σ is constant. Then, the effect on an initial power law spectrum, $\phi = AE^{-\gamma}$, will be simply given by the convolution of ϕ with the error function

$$\tilde{\phi}(E) = \int_0^\infty dE' P(E, E') \phi(E').$$ (2)

In this case the integral can be carried out analytically but is not very illuminating. For high energy E it can be expanded to give

$$\tilde{\phi}(E) = AE^{-\gamma}\left(1 + \gamma(\gamma - 1)\frac{\sigma^2}{2E^2} + \dots\right). \tag{3}$$

At large energies, the effect of the reconstruction energy on the reconstructed flux is very small. This is what is expected for gaussian errors and therefore one can safely neglect this effect. The same occurs for any other initial spectrum. The reconstructed spectrum will differ from the original one by a function which goes to zero rapidly with increasing energy.

Now let's consider the log-gaussian distribution

$$P(E, E') = N\exp(-\frac{\log(E/E')^2}{2\Delta\xi^2}), \tag{4}$$

where N is a normalization constant and $\Delta\xi$ is the standard deviation. Then the convolution of P with the power like initial flux is

$$\tilde{\phi}(E) = \int_0^\infty dE'\phi(E')P(E, E') = AE^{-\gamma}\exp(\frac{\Delta\xi^2}{2}(\gamma^2 - 1)), \tag{5}$$

where we are assuming that $\Delta\xi$ is independent of energy. We see therefore that for a log-gaussian error with a constant standard deviation the effect of the convolution is to change the normalization. This is rather different from the result of Eq.3, where the change was negligible at large enough energies. Here the flux increases (for $\gamma > 1$) by a constant factor. Notice also that the enhancement factor depends strongly on the initial spectrum index, γ. If we put $\gamma = 3$ and $\Delta\xi \sim 15\%$ then $\tilde{\phi}/\phi \sim 1.09$, for $\Delta\xi \sim 30\%$ then $\tilde{\phi}/\phi \sim 1.43$. Alternatively, for $\gamma = 3$, we get a factor two enhancement for $\Delta\xi \sim 0.4$.

In order to see what is the effect of very long tails on the reconstruction of events, let us consider an error function with a long tail

$$P(E, E') = N(\frac{E}{E'})^\alpha(1 + \frac{E}{E'})^{-\alpha-\beta}, \tag{6}$$

where N is a normalization and α and β are constants. The energy is reconstructed with a power law distribution of slope α and $-\beta$ for small and large reconstructed energies. The convolution with the original spectrum gives as before

$$\tilde{\phi}(E) = AK(\alpha, \beta, \gamma)E^{-\gamma}, \tag{7}$$

where K is a constant which depends on the indexes α, β, γ. As in the log-gaussian case the normalization changes but now the effect can be dramatic, depending on the values of α and β. For $\alpha = 4$ and $\beta = 7$ we find $\tilde{\phi}/\phi \sim 1.5$. The log-gaussian distribution can be thought as the limit when α and β are large.

In both cases the normalization is strongly affected. This is easily understood, the long tails on the error functions favor the more numerous low energy events to be reconstructed at higher energy, giving a higher flux. In all these cases we are assuming that there are no systematical energy shifts: the average reconstructed energy for a fixed initial energy

is assumed to be equal to the initial energy. The existence of systematic errors in the reconstruction of energy would worsen considerably our results since in this case the change on the normalization of the spectrum will be linear in $\Delta\xi_{\rm syst}$ rather than quadratic. Also we are assuming that the energy reconstruction function has no strong energy dependence.

Our last example of an error function with long tails is the Moyal distribution which constitutes an approximation to the Landau distribution describing the fluctuations in energy losses of an ionizing particle passing through matter [3].

$$P(E, E') = N \exp(-\frac{1}{2}(\frac{(E - E')}{\sigma} + e^{-\frac{(E-E')}{\sigma}})). \tag{8}$$

We will return to this distribution at the end of the paper when we look in detail at the problem of the GZK cutoff in the UHECR spectrum. Now let's consider an initial spectrum with some feature. In order to keep the analysis simple, let's assume a spectrum with a exponential cutoff.

$$\phi(E) = AE^{-\gamma} \exp(-\frac{E}{E_c}), \tag{9}$$

where E_c is the cutoff energy. The effect of a gaussian error would be again the original flux times a function going rapidly to 1 with energy, *i.e.* the original flux is not modified. However for a log-gaussian error or for a power law error we will have strong modification of the flux. In both cases at low energy ($E \ll E_c$) we will have the previous result, the reconstructed flux is modified by a constant factor. At high energies, $E \gg E_c$, the effect is different. In the log-gaussian error function we have

$$\tilde{\phi}(E) = AE^{-\gamma} \exp(\Delta\xi^2/2(\gamma^2 - 1))e^{-E/\tilde{E}_c}, \tag{10}$$

where now $\tilde{E}_c = E_c \exp(\gamma\Delta\xi^2)$ is the new, smeared, cutoff energy. Again, this effect can be large due to the non linear dependence on γ and $\Delta\xi$; for $\Delta\xi = 0.3$ and $\gamma = 3$ $\tilde{E}_c = 1.3E_c$, for $\Delta\xi = 0.4$ $\tilde{E}_c = 1.62E_c$.

In the case of a power law error function the effect is even more dramatic. At high energy, $E \gg E_c$, the convolution of the error function with the exponential cutoff flux is

$$\tilde{\phi}(E) = AE^{-\beta}E_c^{\beta-\gamma}K'(\alpha, \beta, \gamma), \tag{11}$$

i.e. the cutoff is completely washed out and replaced by a spectrum with a constant slope. The new slope is solely dependent on the error function and not on the original spectrum.

We want now to argue that the non-gaussian error functions given above are realistic. There are several reasons why one expects long tails in the reconstruction of events in cosmic ray experiments. In ground array experiments the energy is estimated by measuring the density of particles in the shower. Usually parameters such as $\rho(600)$, the density of particles at 600 m from the core, are used. It is found that the energy of the shower scales as $E \propto \rho(600)^\kappa$ and $\kappa \sim 1$. But this quantity has non-gaussian uncertainties. First, to calculate $\rho(600)$ (or any similar parameter) the core position and the arrival direction of the shower are needed. Given an arrival direction the $\rho(600)$ is corrected to "zero degrees" arrival direction by an equation like

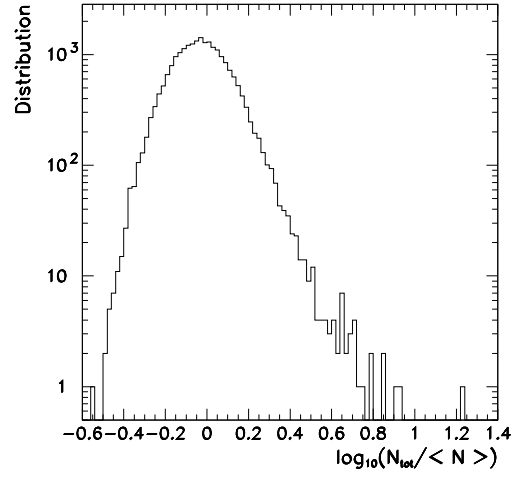$$\rho(600)_0 = \rho(600) \exp(s_0(\sec(\theta) - 1)), \tag{12}$$

Figure 1: Distribution of the total number of charged particles at ground level for protons of $E = 1$ EeV generated using Aires with Sibyll hadronic generator.
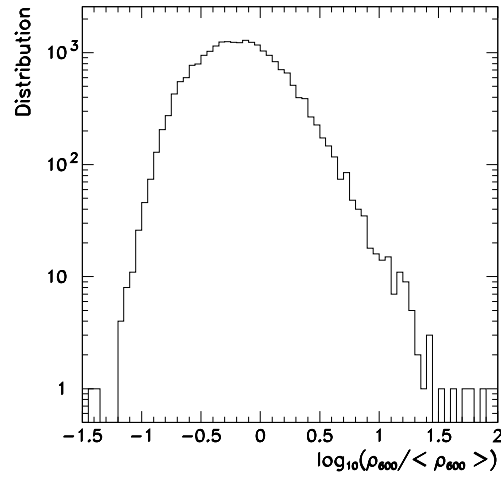

Figure 2: Distribution of $\rho(600)$ at ground for protons of $E = 1$ EeV generated using Aires with Sibyll hadronic generator.

where θ is the zenith angle and s_0 is a constant. Even if the arrival direction is reconstructed with a gaussian probability function the resulting $\rho(600)_0$ is not. However, this effect is, we believe, small at all but the largest zenith angles. The effect on the core reconstruction is more important. The density of particles at a distance r from the shower core can be parameterized by the NKG lateral distribution function

$$\rho(r) = K r^\eta (1 + r/R)^{\eta'}, \tag{13}$$

where K, R, η, η' are constants. Then, the error in the density due to the error in the determination of the core position would be generally non-gaussian.

Finally, any indirect measurement of cosmic rays based on shower development is subjected to shower fluctuations. It is well known that the total number of particles at fixed depth have large non-gaussian fluctuations, the total number of particles r.m.s scales as the energy. The same occurs for $\rho(600)$, since it is related to the total number of particles. In figure 1 we can see the distribution of the total number of particles at ground level in a shower of fixed energy. We can appreciate that the distribution of the number of particles is far from a gaussian and has long tails. In figure 2 we show the distribution of the density at 600 meters for showers of fixed energy. More than 20000 protons showers of energy 10^{18} were simulated with the Aires Monte Carlo code [4]. As before the distribution is non-gaussian with tails extending to more than 1 order of magnitude above the average.

This was known for a long time but rarely used in the analysis of cosmic ray physics [5]. Apparently, it is a generic phenomena in probability theory and does not contradict the central limit theorem [6].

3   The GZK Cutoff

Our results are not directly applicable to any real experiment. Rather, experiments should ascertain what are their error reconstruction functions and take them into account in the calculation of the spectrum. Also, events with large fluctuations could be cut off by other methods, improving therefore the energy reconstruction and avoiding these unwanted effects. Particularly important would be to get rid of any power like tail in the reconstruction error functions, since they have the most devastating effects on the spectrum reconstruction. On the other hand this effect could, at least in part, explain the current experimental situation where the normalization of the cosmic ray flux for different experiments is different and the existence of the GZK cutoff is controversial. Values of $\Delta\xi \sim 30 - 40$ % are usual in cosmic ray physics and as said before with such values there are important effects on the shape and normalization of the resulting spectra.
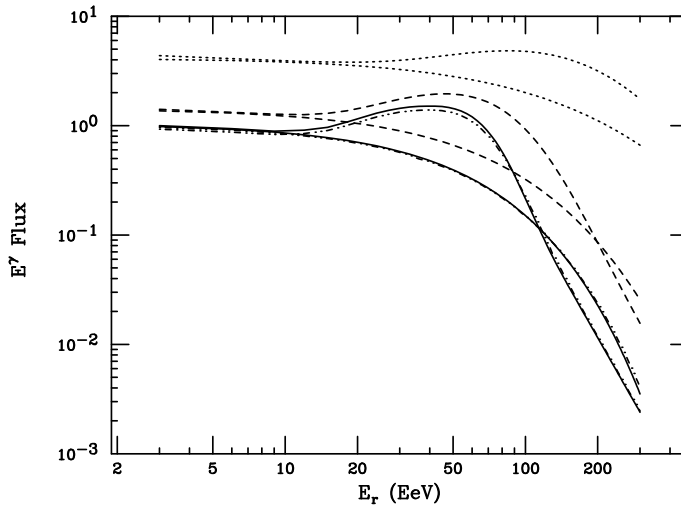
Figure 3: Expected flux for log-gaussian energy reconstruction error function with $\Delta\xi = 0.1, 0.3, 0.6$ (continuous, dashed, and dotted lines) for a flux with exponential cut-off at $E = 50$ EeV and for a cosmological distribution of sources.

The current experimental situation is aggravated by the low statistics in the high energy part of the spectrum. The Agasa experiment has measured 58 [8] events above 4×10^{19} eV, of them 8 are above 10^{20} eV. Using a realistic spectrum obtained from an uniform cosmological distribution of sources [9] and using a log-gaussian error function with $\Delta\xi = 0.4$ we obtain that for a total of 58 events above 4×10^{19} eV, the probability of having 8 or more events above 10^{20} eV is 2%. The average number of expected events would be 3. In figure 3 the convolution of different spectrum with energy error functions is shown. In the case of a $\Delta\xi = 0.3$ and the cosmological distribution of sources of Teshima and Yoshida [9] one can see that at $E = 200$ EeV only a 0.1 reduction on the flux is expected.

The number of AGASA events above the GZK cutoff seems to establish its absence. However, if we introduce non-gaussian errors, we stress once again, the observation of such events becomes compatible with the GZK cutoff. To illustrate quantitatively the possibility of not seeing the GZK cutoff because of non-gaussian errors in the energy determination we consider the Moyal distribution of equation 4 (with approximately the same half-width of a gaussian distribution of a given variance, as seen in fig. 4) to each energy sampled from the AGASA spectrum with a exponential cutoff, that is a power law spectrum with an index of 2.78 multiplied by an exponentially falling spectrum starting at 4×10^{19} eV. These errors alter the shape of the original spectrum and the expected number of events above the GZK cutoff . We summarize the results of this analysis in fig. 5. We can compare the probability of finding events above GZK for measurements affected by gaussian errors (10%, 20%, and 30%) with measurements affected by Moyal-distributed errors (with the same half-width of gaussians with standard deviation of 10%, 20%, and 30% of energy). We see that considering only gaussian errors we cannot explain the observed number of events above GZK by AGASA (indicated by an arrow in figure 5) and would be led to conclude that there is a cutoff in
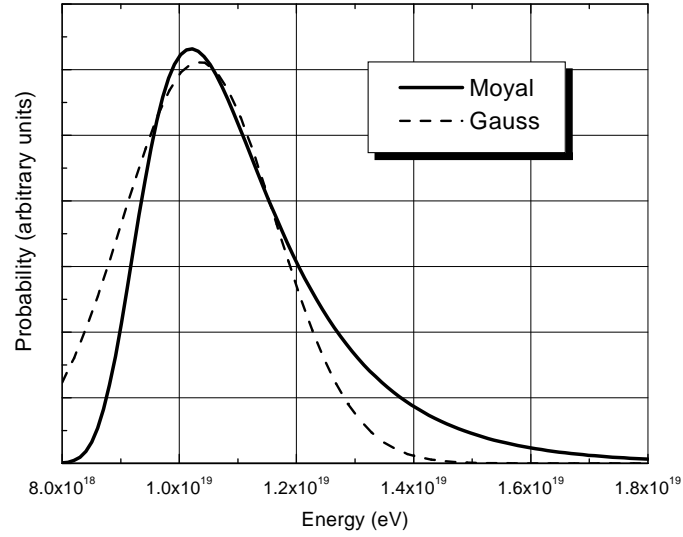
6

Figure 4: The Moyal distribution of Eq. 4 compared to a gaussian distribution having approximately the same half width
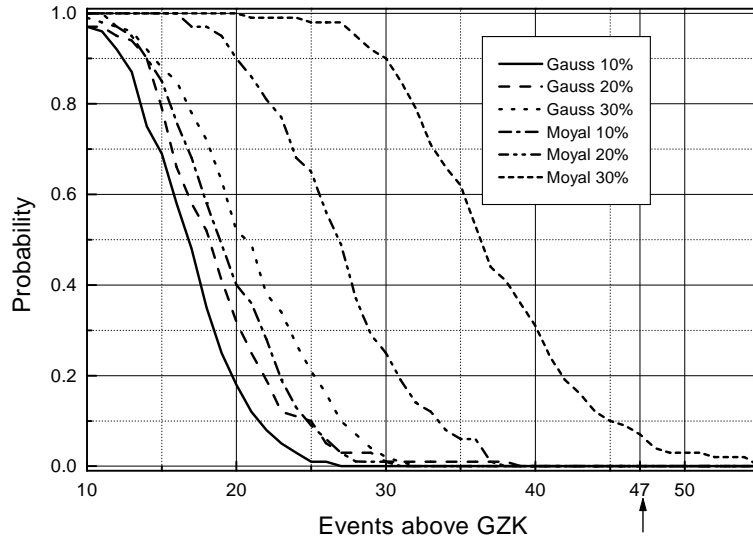

Figure 5: Probability of obtaining events above the GZK cutoff for different resolutions (error functions)

the UHECR spectrum. However, for Moyal-distributed errors (30% of energy) we have an appreciable probability of finding more than 47 events above the GZK cutoff. We have shown in this paper that the experimental resolution curve when deviating from a gaussian distribution severely impacts the shape and normalization of a steeply falling spectra near its end point. Spectral features such as a cutoff may disappear from the reconstructed spectrum due to the effect of long tails in the resolution(error)function of the experiment.

References

[1] E. Holzschuh, *Rep. Prog. in Phys.* **55** (92) 1035; E.W. Otten, *Prog. Part. Nucl. Phys.* **32** (94) 153.

[2] M. Nagano and A.A. Watson, *Rev. Mod. Phys.* **72** 689 (2000).

[3] J. E. Moyal, *Phil. Mag.* **45** (1955) 263.

[4] AIRES: A System for Air Shower Simulation, S.J. Sciutto, *Proc. of the XXVI Int. Cosmic Ray Conf.* Salt Lake City (1999), vol. 1, p. 411-414; S.J. Sciutto *preprint archive* astro-ph/9911331.

[5] G. Parente, R.A. Vázquez, and E. Zas, Proceedings of 24th the ICRC, Vol.1 p.377 (1995).

[6] M.I. Tribelsky, preprint math.PR/0106037, June 2001.

[7] M. Ave, *private communication*

[8] N. Hayashida *et al*, *Astrophys.J.* **522** (1999) 225.

[9] S. Yoshida and M. Teshima, *Prog. Theor. Phys.* **89** (1993) 833.